doi:10.1017/S0003055409090224

Moral Bias in Large Elections: Theory and Experimental Evidence

TIMOTHY FEDDERSEN *Northwestern University*
SEAN GAILMARD *University of California at Berkeley*
ALVARO SANDRONI *University of Pennsylvania*

We argue that large elections may exhibit a moral bias *(i.e., conditional on the distribution of preferences within the electorate, alternatives understood by voters to be morally superior are more likely to win in large elections than in small ones). This bias can result from* ethical expressive *preferences, which include a payoff voters obtain from taking an action they believe to be ethical. In large elections, pivot probability is small, so expressive preferences become more important relative to material self-interest. Ethical expressive preferences can have a disproportionate impact on results in large elections for two reasons. As pivot probability declines, ethical expressive motivations make agents more likely to vote on the basis of ethical considerations than on the basis of narrow self-interest, and the set of agents who choose to vote increasingly consist of agents with large ethical expressive payoffs. We provide experimental evidence that is consistent with the hypothesis of moral bias.*

In this article, we provide evidence that voters in large elections tend to vote against their material self-interest and to vote for a morally or ethically appealing alternative. It may seem puzzling that voters might behave differently in small elections than they do in large ones. However, we show that such behavior is a logical consequence of voters having a conflict between obtaining a better material outcome and choosing a moral action. We develop a simple model of this conflict and show that decreasing the probability that a single vote is decisive (i.e., pivot probability) reduces the importance of outcomes relative to actions in voter decision making. Because pivot probability is generally small in large elections, alternatives that are understood by voters to be morally superior are more likely to win in large elections than in small ones. Thus, compared to the preferences of voters, election results will be biased in favor of moral alternatives.[1] The model produces a set of predictions that we test in a laboratory experiment.

To clarify ideas, consider a situation in which two outcomes are possible: A and B. Assume that B gives higher material benefit than A to each voter in the electorate, but all agree that A is morally superior. This might be the case if, for example, B gives high monetary returns to all voters and imposes high costs on a population of nonvoters, whereas A gives moderate benefits to voters and nonvoters alike. Suppose that each voter, if given the choice between outcome A and outcome B, will choose B. In standard decision theory, this is equivalent to saying that this person *instrumentally prefers* B to A. However, in an election (without abstention), no voter decides the outcome unilaterally. Rather, each voter has two actions: vote for A and vote for B such that voting for A leads to higher probability of outcome A than voting for B. The increased probability of outcome A achieved by voting for it is called the *pivot probability* and denoted by $p > 0$. If voters only have instrumental preferences, they will vote for B (independent of the magnitude of the pivot probability) because that increases the probability of the instrumentally preferred outcome.

Standard theory assumes that preferences over outcomes induce preferences over actions. The logic of moral bias rests on the idea that people have a desire to act ethically independent of the outcome produced. To model this idea, we assume that voters obtain a small positive payoff by the act of voting for A. That is, voters have an *expressive preference* for the act of voting for A because by voting for A they vote for the alternative that is, by assumption, morally superior. As pivot probabilities decrease, the difference in instrumental payoffs between voting for A and B diminish, whereas the expressive payoff remains constant.[2] A voter with both instrumental and expressive preferences chooses as follows. If the pivot probability is large, then instrumental payoffs dominate the expressive payoffs and the voter chooses to vote for B. However, when the pivot probability is sufficiently small, expressive payoffs dominate and the voter chooses to vote for A.

Timothy Feddersen is the Wendell Hobbs Professor of Managerial Economics, Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208 (tfed@kellogg.northwestern.edu).

Sean Gailmard is Assistant Professor, Charles and Louise Travers Department of Political Science, University of California at Berkeley, 210 Barrows Hall, Berkeley, CA 94720 (gailmard@berkeley.edu).

Alvaro Sandroni is Professor, Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104 (sandroni@econ.upenn.edu).

The authors thank Thomas Palfrey, Jasjeet Sekhon, Laura Stoker, and Rob Van Houweling for helpful comments; seminar participants at Dartmouth, Essex, Florida State, Oxford, Penn, and UC Berkeley; panelists at the 2006 annual meeting of the Midwest Political Science Association and the 2007 Society for the Promotion of Economic Theory conference; and Jenny Cheng for research assistance.

[1] The term "bias" here simply refers to a systematic departure from the outcome that would be obtained if all voters acted instrumentally based on their preferences over outcomes. It is a bias against an implicit baseline of selfish instrumental voting, a choice we make given the prominence of this model in the literature since at least Downs (1957). It should also be noted that we do not advance an argument about whether moral bias is normatively "good" or "bad," in general.

[2] In fact, all that is really necessary is that the expressive payoff for acting morally diminishes with pivot probabilities at a slower rate than the instrumental payoffs.

We call this the *preference effect* of a change in pivot probability.

Small pivot probabilities are a fundamental characteristic of large elections. Hence, an expressive preference for voting for a moral outcome generates a *moral bias*[3]: controlling for the distribution of preferences within the electorate, alternatives that are understood by voters to be morally superior are more likely to win in large elections than in small ones, even if they are contrary to voters' material self-interest.[4]

When there is no cost to vote, the impact of expressive preferences on election outcomes depends entirely on the percentage of voters with expressive preferences. If such voters are a small fraction of the electorate, their impact may be minimal because many others would vote for *B*. However, when small costs to vote are introduced, the impact of expressive preferences is greatly amplified. This follows because voters without an expressive payoff become more likely to abstain as pivot probability declines. Voters with expressive preferences, however, continue to vote and become more likely to vote for *A* as pivot probabilities decrease. This follows because such voters obtain an expressive payoff by voting for *A* that does not decline with pivot probability. Thus, in sufficiently large elections when voting is costly, we may expect the electorate to consist entirely of voters with expressive preferences who vote for *A*. We call this a *turnout effect*. Thus, moral bias has two causes in principle, a preference and a turnout effect.

Beyond clarifying the basic mechanics of moral bias, the contribution of this article is to provide experimental evidence for the preference and turnout effects at the level of individual decision makers. We construct an experiment in which a group chooses between one of two alternatives, *A* and *B*. The group is subdivided into *A* types who obtain a high payment under outcome *A* but nothing under *B*, and *B* types who obtain a high payment under outcome *B* but a smaller (although relatively nontrivial) payment under *A*. The majority of each group is *A* types. Therefore, alternative *A* maximizes the sum of payments, gives nearly equal payments to everyone, and maximizes the minimum payment. For these reasons, we call *A* the *ethical* alternative.

In experimental elections, analyzing the effects of changes in pivot probability is complicated because of a multiplicity of equilibria.[5] To avoid this problem, we simulate an election with costly voting by using a decision mechanism in which pivot probability is controlled directly as a treatment variable. In our experiment, a subset of *B* types is designated as *active* and may either vote for *A* or *B* at a cost ($c > 0$), or abstain at no cost. The outcome is determined when one active *B* type is selected at random. If the selected individual has not abstained, his or her vote determines the outcome. If the selected individual has abstained, then *A* and *B* are chosen with equal probability. The number of active individuals therefore precisely determines the probability that an active individual's vote is pivotal.[6] Because only *B* types can vote, any vote for *A* necessarily runs counter to the voter's material interest.

Our experimental results show that remarkably small electorates are nevertheless large enough to generate behavior that is consistent with the moral bias hypothesis. Varying pivot probability from $\frac{1}{11}$ to 1, we find that selfish votes (i.e., votes for *B*) are strongly positively related to pivot probability, but ethical votes (i.e., votes for *A*) are either insensitive or negatively related to pivot probability. Moreover, the ethical alternative *A* is significantly more likely to be the collective choice as pivot probability declines.

In most elections, pivot probabilities are likely to be significantly below $\frac{1}{11}$. Hence, our experimental results are consistent with the hypothesis of a moral bias in large elections. The point here is not that elections in more populous states (e.g., California) are more likely to be organized around ethical concerns than elections in less populous states (e.g., South Dakota): pivot probabilities are likely to be miniscule in both. Rather, the point is that electioneering and discussions of vote choice in any mass election are likely to emphasize moral values. "Values voting," sociotropic motivations, and other departures from material self-interest are to be expected with rational expressive voters even though the same voters may well act on the basis of material self-interest in their everyday life.

The remainder of the paper proceeds as follows: we present a brief literature review; formally define instrumental and expressive utilities for voters and derive hypotheses about behavior as a function of pivot probability; describe the experiment design to test these hypotheses; present results from the experiment; and conclude.

RELATED LITERATURE

Voting Models

Our model is part of a long line of noninstrumental or partially instrumental models of turnout and vote choice. It is well known that the standard rational choice model has trouble explaining turnout in large elections with costs to vote.[7] Riker and Ordeshook (1968), Tullock (1971), Brennan and Buchanan (1984), Brennan and Lomasky (1993), Scheussler (2000), and Feddersen and Sandroni (2006a, 2006b) have proposed *expressive* theories of voting to explain turnout in large

[3] We adopt the term "moral bias" because our basic theory is motivated by ideas of morality. Our experimental design allows us to observe whether agents act in a manner that is consistent with our theory, but we cannot observe agents' actual motivations.

[4] Note that standard theory allows for voters whose instrumental preferences incorporate altruistic, inequality-averse, inefficiency-averse, and other types of moral concerns as well. However, voters with only instrumental preferences will not change their behavior as a consequence of a change in pivot probabilities. Hence, instrumental moral preferences will not produce a moral bias.

[5] See Levine and Palfrey (2007) and Duffy and Tavits (2006).

[6] See Grether and Plott (1979) for an example of such an approach.

[7] See Levine and Palfrey (2007) for experimental evidence that standard models can explain turnout in large elections.

elections.[8] In expressive voting models, individuals are motivated to vote not out of a desire to directly impact the election outcome, but out of a sense of civic obligation or a desire to (seem to) act ethically by supporting morally appealing causes or candidates. In the Riker-Ordeshook model, this creates an extra payoff for any vote choice other than abstention; the others essentially modify the Riker-Ordeshook model by attaching the d term to a specific vote choice.[9]

Although the focus in this article is moral bias, it should be noted that other kinds of biases have been discussed in the literature as well. Clearly, anything that may lead the preferences of the electorate to differ substantially from the preferences of the population as a whole can lead to nonrepresentative outcomes. For example, costs to vote are known to significantly decrease turnout (see Levine and Palfrey 2007; Riker and Ordeshook 1968) and may bias election results in favor of those with lower costs to vote. In small elections with high pivot probabilities, even voters with relatively high costs may participate, whereas in large elections they will abstain. If costs are not distributed uniformly throughout the population, outcomes in large elections may be systematically different from the outcomes that would occur in small elections. Asymmetric information can also produce systematic differences between outcomes in small and large elections. For example, Feddersen and Pesendorfer (1996) demonstrate that election outcomes will tend to approach those generated by an electorate of fully informed voters as the size of the electorate grows.

Laboratory Experiments

A small number of experimental papers have analyzed expressive voting, with mixed results. Tyran (2005) finds essentially no support for expressive voting. There are several salient differences between our experiment and Tyran's that could account for the divergent results. First, in Tyran's experiment, the "ethical" choice requires subjects to forfeit their entire endowment (granted by the experimenter); this is much more selfishly costly than in our design. This has two effects: first, conditional on pivot probability, it raises the selfish cost of voting ethically. Second, it may affect whether subjects regard the "ethical" choice as, in fact, more ethical. Self-serving perceptions of fairness may allow subjects to convince themselves that they have no duty to give away their own endowment as a pure transfer (Babcock and Lowenstein 1997; Messick and Sentis 1979). In addition, Tyran does not control pivot probability; instead, he elicits subject beliefs about it for one treatment and, from this, infers what "reasonable" beliefs would be in other treatments. It is possible that subjects do not accurately perceive their pivot probabilities, or that their subjective perceptions in cases where pivot probability is not elicited by Tyran do not match his inference. In contrast, because we control pivot probability directly and demonstrate it to subjects, our design more directly links pivot probability and behavior.

In an experiment on the effect of inequality aversion on voting for redistribution, Tyran and Sausgruber (2006) include a brief check for expressive voting. They elicit pivot probability estimates from voters in groups of five, and find that voters who believe they are pivotal vote less frequently for costly redistribution than voters who believe they are not pivotal, although the difference is not statistically significant.

Fischer (1995) obtains stronger support for expressive voting, but with relatively little data to identify the effect of pivot probability (20 subjects in four treatments). Moreover, pivot probability is not controlled directly, nor are subject beliefs about it elicited. Therefore, as in Tyran (2005), the link between pivot probability and behavior remains somewhat circumstantial in the data. Fisher's experiment is a follow-up to Carter and Guerette (1992), where subjects voted either to keep an entire endowment (either \$6 or \$9) or receive nothing and donate a smaller amount (\$2) to charity. Carter and Guerette more directly controlled pivot probability: each subject registered a choice for how his or her own endowment would be used, and with a probability specified by the experimenters, this choice was in fact implemented for that subject; otherwise, the choice was made for that subject by a coin flip. Carter and Guerette uncover modest support for the expressive voting hypothesis in one sample of subjects (the effect of pivot probability on votes for charity giving is significant at the .10 level in a 1-tailed test), but found it sensitive to the selfish cost of charitable giving (based on another sample where subjects forfeited \$9 rather than \$6). As Carter and Guerette note, the option of charitable giving is inefficient in their experiment, in addition to being very costly for subjects; both factors may mitigate the ethical appeal of this option to the subjects.

Our theory and design highlight several features that are not accounted for in these papers taken as a group. First, none allow abstention or make voting per se costly. As we show in the formal model, abstention and costly voting interact with expressive preferences to determine both individual and group choices. Second, we find an effect of pivot probability not only on individual decisions to vote selfishly, but also on collective outcomes. These findings suggest that moral bias may be "economically relevant" in large elections. In addition to allowing us to address these issues not considered in previous experimental research, our design imposes smaller costs on subjects for voting against their material self-interest and implements direct control of pivot probability in a setting where subjects distribute money among themselves. As our theory suggests, smaller costs of voting against one's self-interest can affect both the price and frequency of nonselfish voting. It may also affect whether subjects perceive

[8] Uhlaner (1989) and Morton (1991) also explore noninstrumental theories of turnout but focus on group mobilization instead of expressive benefits. See also Harsanyi (1977) and Coate and Conlin (2004).

[9] This can also be thought of as a "warm glow" payoff (Andreoni 1990). See Andreoni (2006) for a review of the literature on warm glow giving.

nonselfish voting as ethical in the first place. Although it is informative to demonstrate that nonselfish voting can be extinguished if it is costly enough, we believe it is also useful to determine whether expressive preferences affect collective outcomes, even if they can be dominated by selfish preferences.

A different literature has addressed the valuations that individuals reveal in surveys for resources not traded in markets (e.g., national parks), with results that are suggestive of ethical expressive preferences. In a common survey technique known as contingent valuation, respondents are asked how much they would be willing to pay to secure some nontraded resource, in the hypothetical situation where they must actually pay that amount. "Hypothetical bias" occurs when these elicited values are greater than values that individuals would reveal in nonhypothetical situations (Cummings, Harrison, and Rutström 1995; Harrison and Rutström 2008). A meta-analysis of experiments on contingent valuation found that individuals overstate actual valuations by up to a factor of three in hypothetical choice situations (List and Gallet 2001). A hypothetical choice is one in which the stated preference by definition cannot affect the outcome. In the context of our model, this would be equivalent to setting the pivot probability to 0.

Our work is also related to the large literature on social preferences (e.g., altruism, inequality aversion) and giving in the Ultimatum and Dictator games (see Camerer 2003, for a partial review). In the two-player Dictator game, as in our results (where subjects are randomly chosen as dictator in a multilateral setting), subjects tend to sacrifice to improve recipients' payoffs. Previous work (Bolton and Ockenfels 2000; Fehr and Schmidt 1999) has suggested that this is because such sacrifice reduces inequality in payoffs, which subjects dislike; Charness and Rabin (2002) note that it also increases the minimum payoff and this may have more explanatory power. However, the social preferences literature does not directly inform the expressive versus instrumental view of preferences; social preferences could underlie both ethical expressive preferences, and ethical instrumental preferences. Jankowski (2002), Fowler (2006), Tyran and Sausgruber (2006), and Edlin, Gelman, and Kaplan (2007) all incorporate social preferences into models of instrumental voting. Such motivations can rationalize voting nonselfishly in an instrumental context (even when abstention is possible and voting is costly); Fowler finds that individuals who are more generous in the Dictator game are more likely to vote in U.S. elections, and Tyran and Sausgruber (2006) find that inequality aversion affects voting in a laboratory experiment on redistribution. However, these motivations still imply that participation decreases as pivot probability declines, as we explain further in the next section. This fact allows us to distinguish the effects of ethical expressive preferences from those of ethical instrumental preferences.

Finally, our design relates to the experimental literature on redistribution behind the "veil of ignorance" (Frohlich and Oppenheimer 1992). This is operationalized (see, e.g., Beckman et al. 2002) by having players allocate shares of a resource among unspecified members of a group, and then randomly assigning them to a position in the group. This is similar to our design in which players specify monetary payoffs to group members, and the actual distribution implemented is randomly chosen. However, in veil of ignorance designs, it is typically possible for a player specifying a lopsided distribution to then obtain small payoffs under that distribution. This is not possible in our design. As a result, individuals cannot ensure themselves in our design by voting for the ethical alternative.

Survey Research

Our findings dovetail with the large literature exploring the ability of self-interest to explain political attitudes and vote choice expressed in surveys. In general, this evidence indicates that self-interest explanations in general and "pocketbook" considerations in particular only weakly account for attitudes and vote choice in the American electorate (see Sears and Funk 1991, for a review). Consider two classic examples: Kinder and Kiewiet (1979) show that at the level of individual voters, material self-interest does not account well for voting in U.S. congressional races; and Sears and Citrin (1985) show that symbolic attitudes about government in general and government spending on welfare affected support for property tax reduction referendums in California. However, Sears and Citrin also show that with high material stakes for homeowners, self-interest also clearly affected support for the tax changes.

FORMAL MODEL

Consider a group of $n > 0$ individuals that must choose between two options, A and B. The group is composed of two subgroups, A types who get a higher monetary reward if option A is the outcome and B types who get a higher monetary reward when option B is the outcome. Let $n_A > 0$ and $n_B > 0$ denote the number of individuals of each type where $n_A + n_B = n$.

The set of B types is further subdivided into *active* and *inactive* individuals. Let n_β be the number of active B types. Only active B types have a chance to influence the group decision.

Active B types simultaneously and privately choose one of three options: abstain, vote for A, or vote for B. The group decision is determined by selecting one active B individual at random. If the selected individual has voted, then his or her vote determines the outcome. If he or she has abstained, then the group outcome is determined by the flip of a fair coin.

Monetary rewards are provided in Table 1.

The term c and x are parameters in the model where $0 < c < 1$ corresponds to a monetary cost of voting. The parameter x corresponds to a monetary premium for B types if option B is the outcome. A types receive a monetary reward of $1 - c$ if alternative A wins the election, and 0 otherwise.

We assume that $\frac{1}{2} > x > 2c > 0$ and $n_A > n_B$. These assumptions ensure that alternative A minimizes

TABLE 1. Monetary Rewards under Options A and B

	A Type	Active *B* Type Who Vote	Other *B* Types
Option *A*	$1-c$	$1-c$	1
Option *B*	0	$1+x-c$	$1+x$

inequality in terms of monetary rewards, maximizes the sum of monetary rewards, maximizes the minimum reward, and gives a higher monetary reward to a majority of the group. For these reasons, we say that A is the *ethical outcome*.[10] Loading these ethical concerns onto one alternative makes it impossible to specify exactly which ethical motivations are behind our results. This issue is beyond the scope of our analysis because we address the effects of expressive voting for ethical alternatives, not the secondary question of which particular ethical considerations carry the greatest weight.

Payoffs

The payoffs for each choice are modeled as follows. The payoff of voting for alternative A is

$$\pi_A(n_\beta, x) = \frac{1}{n_\beta}(1+\delta) + \left(1 - \frac{1}{n_\beta}\right) \times (1 + \delta + q^*(x-\delta)) - c + d - \varepsilon_A.$$

The payoff of voting for alternative B is

$$\pi_B(n_\beta, x) = \frac{1}{n_\beta}(1+x) + \left(1 - \frac{1}{n_\beta}\right) \times (1 + \delta + q^*(x-\delta)) - c - \varepsilon_B.$$

The payoff of not voting is

$$\pi_\phi(n_\beta, x) = \frac{1}{n_\beta}\left(1 + \frac{x+\delta}{2}\right) + \left(1 - \frac{1}{n_\beta}\right) \times (1 + \delta + q^*(x-\delta)) - \varepsilon_N.$$

The variable q^* is the probability that alternative B is selected when the decision maker's vote is not pivotal. The term ε_k is a stochastic payoff disturbance with $E(\varepsilon_k) = 0$ and $Var(\varepsilon_k) = v$ for $k \in \{A, B, N\}$ (e.g., ε_k may be assumed to follow a Type I extreme value distribution independent of $\varepsilon_j, j \neq k$ as in multinomial logit models). This term anticipates that choices in experiments typically vary, even for a given individual with all parameters held fixed. As we explain, n_β, x, and c are all controlled in the experiment, with x and c fixed as in Table 1 and n_β varying in the experiment.

Our payoff functions not only include monetary payoffs, but also allow for the possibility of two different kinds of subjective payoffs—an *ethical instrumental payoff* $\delta > 0$ when the alternative A is the outcome, and an *ethical expressive payoff* d obtained as a consequence of voting for alternative A. Both types of subjective payoffs may be behaviorally important. In the case of the payoff δ, instrumental voters may depart from selfish behavior if they take into account the monetary rewards of others.[11] The payoff d captures the expressive rather than instrumental motivation for supporting A. We call these "ethical expressive" motivations and model them with a payoff $d > c$ that voters obtain by voting for option A.[12] Given these payoffs, we can now define the comparative statics of the general model; Appendix A reviews comparative statics for special cases, including selfish instrumental voting.

Conditional on voting, a voter is expected to vote for A over B if $E(\pi_A(n_\beta, x)) \geq E(\pi_B(n_\beta, x))$, which with some algebra reduces to the following:

$$d \geq \frac{x-\delta}{n_\beta}.$$

Note that a voter who receives neither type of ethical payoff (i.e., $d = \delta = 0$) will always vote for B rather than A. Moreover, voters who obtain no ethical *expressive* payoff ($d = 0$) will not vote for A unless the ethical instrumental payoff is greater than the material gain that results from alternative B (i.e., $\delta > x$). In contrast, for sufficiently small pivot probability ($1/n_\beta$), a voter with even a small expressive payoff will vote for A rather than B.

A voter is expected to vote for A rather than abstain if

$$d - c \geq \frac{x-\delta}{2n_\beta}.$$

In this case, a voter who receives no ethical instrumental or expressive payoffs will always abstain rather than vote for A. A voter who receives no expressive payoff ($d = 0$) will only vote for A if δ is sufficiently large. However, as pivot probabilities decrease, such a voter will ultimately abstain. In contrast, a voter who receives a sufficiently large expressive payoff ($d > c$) will always turn out and vote for A if pivot probabilities are sufficiently small.

Finally, a voter is expected to vote for B over abstention if

$$\frac{x-\delta}{2n_\beta} \geq c.$$

As pivot probabilities decrease, no voter is expected to vote for B.

[10] In fact, the assumption $x > 2c$ could be replaced by the weaker assumption that $x > c$. However, this stronger condition simplifies the exposition and is consistent with the monetary payoffs we offered in our experiments. We discuss the case $2c > x > c$ in Appendix A.

[11] This could capture, for example, distributional concerns or difference aversion; see the distributional utility function in Charness and Rabin (2002) for a general representation.

[12] The assumptions that $d > c$ and that A is the ethical option (for which the expressive payoff d occurs) ensure that the behavior of the ethical expressive voters is different qualitatively from the behavior of the selfish voters. We relax these assumptions in Appendix A.

The presence of an ethical expressive component to payoffs produces much different behavior than an ethical instrumental payoff alone (whether that ethical instrumental payoff δ is small or large). The ethical expressive payoff has an important effect on the relationship between pivot probability and the collective choice in the election. Specifically, voters may exhibit a propensity to vote for the alternative B when pivot probabilities are high and a propensity to vote for the alternative A when pivot probabilities are low. This is the *preference effect* of pivot probability alluded to at the beginning of the article. This may seem counterintuitive, but it has a straightforward intuition. As pivot probabilities decrease, the choice of which candidate to vote for becomes essentially hypothetical because it does not have much impact on the voter's material payoff. Therefore, the potential benefit from voting selfishly becomes small, whereas the expressive payoff from voting for the ethical alternative, which is not affected by pivot probability, stays constant.

A second behavioral difference for voters with ethical expressive payoffs and instrumental payoffs is that in the former case the incentive to vote may be nondecreasing or even increasing as pivot probabilities decrease, whereas in the latter cases the incentive to vote is decreasing as pivot probabilities decrease. This is the *turnout effect* of pivot probability alluded to previously.

Hypotheses

From this analysis, one can predict the effects of pivot probability on participation and vote choice by agents, and therefore, the effects of pivot probability on the probability that each option is chosen for the group.

Specifically, suppose all agents have payoffs as specified in the previous model. Then, if $d > c$, the model predicts that as pivot probability increases,

1. The probability an agent votes for B rather than abstains is nondecreasing. This is because instrumental motivations behind selfish voting increase with pivot probability.
2. The probability an agent votes for A rather than abstains is nonincreasing. This is because of ethical expressive motivations to support option A. These motivations do not result from the possible effect of a vote on the election outcome, so are not sensitive to decreasing pivot probability. However, this motivation to vote against A is decreasing in pivot probability.
3. Conditional on voting, the probability an agent votes for A rather than B is nonincreasing. This follows from the preference effect. When pivot probability increases, vote choice is further from a hypothetical choice, so the noninstrumental component of utility has less weight.
4. The probability the group selects the ethical option A is nonincreasing. This follows from both the preference and turnout effects.

EXPERIMENT DESIGN

We test the predictions from the section previous in a laboratory experiment. The experiment implements the decision model of an election. For experimental purposes, the decision model of an election has advantages over a game-theoretic election model with endogenous pivot probability. Our approach allows us to directly manipulate pivot probability, the key causal variable in our theory, as a treatment variable, and thus ensure that it is independent of individual tendencies to weigh ethical considerations versus selfish payoffs in making decisions.

Our experimental design also allows us to induce specific monetary rewards for the options facing the group. Even though we cannot fully control preferences, control over monetary payoffs still allows us to determine which option is materially beneficial to voters and which option is ethical in several respects.

Our decision process is relevant to elections in a strict sense because we precisely capture the decision-relevant consequences of being pivotal in an election. In particular, in our decision mechanism as in an election, in the event that one's vote would decide the outcome in favor of one alternative or another, the decision to abstain creates a tie, whereas the decision to vote is equivalent to deciding the group choice. However, it must be noted that our decision process differs from an election in the responsibility an individual may perceive over the outcome. In our decision process, an individual knows whether he or she was the dictator ex post and therefore unilaterally responsible for the outcome. In an election, one's vote is pivotal only if all other votes result in a tie. Although the pivotal voter is by definition decisive, this decisiveness is conditioned by the choices of other voters as well. Thus, the pivotal voter is not unilaterally responsible for the outcome. Thus, compared to our decision process, an election may dampen an individual's sense of responsibility for the outcome, although not completely eliminate it.

A few caveats are in order here regarding the relationship between our theory of moral bias and our experimental design. Our design captures the central decision-theoretic problem facing voters in an election: whether to vote for a candidate or abstain as a function of the probability one's action influences the outcome. However, in contrast to an election, in our setting an individual knows if he or she was decisive ex post and therefore unilaterally responsible for the outcome. In an election, one's vote is pivotal only if all other votes produce a tie. Hence, no single individual is ever unilaterally responsible for the outcome of an election even ex post. Thus, compared to our decision process, an election may dampen an individual's sense of responsibility for the outcome. It would be worthwhile to perform experiments similar to ours but in the context of experimental elections.

We should also note that our experimental design does not allow us to conclude that voters have moral motivations. The fact that a voter has chosen an alternative that (at least to us) seems morally appealing does not allow us to conclude that experimental subjects were motivated by moral considerations. More research is required to support the empirical claim that voters are morally motivated.

In our experiment, subjects were divided into groups and offered monetary incentives, as specified in Table 1.[13] The experiment consisted of a sequence of rounds in which groups chose option *A* or option *B*. A round, in turn, consists of four stages. In stage 1, a group of *n* subjects (where $n \leq$ the number of subjects in the session) is randomly partitioned into two subsets with n_A *A* members and n_B *B* members. Furthermore, n_β of the *B* types are randomly designated as *active B types*. Each subject in a group is informed of the number of people of each type before any decisions are made. Subjects know which category they themselves are in but are not informed of the identity of other individuals in these categories. A *B* type learns whether he or she is an active type before making any decisions.

In stage 2, each active *B* type must choose whether to vote. If he or she chooses to vote, then he or she pays a small cost *c* and specifies one of the two outcomes *A* or *B*. All other subjects (i.e., all *A* types and inactive *B* types) have no decision to make.

In stage 3, after all active *B* types make their participation and vote choices, one active *B* subject is randomly selected from the set of all active subjects. The probability that a given active subject is selected is $\frac{1}{n_\beta}$. Note that any active subject can be randomly selected at this stage, regardless of whether they have chosen to vote.

Stage 4 determines the group choice. If the active subject selected at stage 3 has chosen to vote, then the outcome that subject specified at stage 2 is the group choice. If this subject has not voted, then the outcome, *A* or *B*, is chosen by a fair coin toss. Thus, n_β fully determines pivot probability. Note that the decision problem for subjects with preferences as outlined previously is identical whether pivot probability is endogenous or exogenous. Thus, our design maintains the same incentive effect of pivot probability on voting that exists in an election, while allowing for control of this key variable.

At the end of each round, subjects are informed of the group choice, their payoff, whether the decision was made randomly or by an active *B* type, and for subjects who were active *B* types, their own decision and whether they were pivotal. For each round, this information and a subject's type (*A* or *B*) in the round were displayed in a History panel visible on subjects' computer terminals throughout the experiment.

This sequence of four stages makes up a single round of a session of the experiment. After one round is completed, another begins with a new random allocation of *A* and *B* types to groups and a new random draw of n_β active *B* types in each group. A sequence of rounds with groups drawn from a set of participants comprises a session of the experiment.

A triple (n_A, n_B, n_β) is a distinct *treatment* in the experiment; these variables are subject to experimental control. As noted, n_β determines the probability that a vote is pivotal, whereas changes in n_A and n_B determine the collective benefits that result from each outcome, as well as the degree of inequality in the group under option *B*. Treatments were run in *blocks* of 10 to 20 consecutive rounds with the same values of (n_A, n_B, n_β). A subject's type (*A* or *B*) was fixed for all rounds in a given treatment block and randomly redrawn in the next treatment block, whereas Active/Passive status for *B* types changed randomly from round to round within a treatment block. Subject types (*A* or *B*) were fixed within a treatment block to limit repeated play effects from artificially producing ethical votes. Without this feature, round *t* voters might have an incentive to vote ethically in hopes of priming round $t + 1$ voters to do so as well, in case round *t* voters were to switch types. Subjects were informed of these conditions in the instructions and they were demonstrated in practice rounds, and subjects were also informed verbally when a treatment block had ended during the session.

As noted, monetary payoffs in the experiment are determined as in Table 1. In all rounds of the experiment, $c = .10$ denotes the participation cost and $x = .25$ denotes the premium that *B* types earn from option *B* over option *A*. Participants are informed of these parameters in the instruction period and in a table visible to them throughout the experiment.

In the actual experiment, we described the decision situation to subjects in neutral, abstract terms. In particular, we never used the words "selfish" or "ethical" in the experiment. In addition, we referred to active types as *active* and to those who decided to vote as subjects who choose to be *available.* This mitigates a potential contaminating effect of "tipping off" the subjects about the kind of behavior that is somehow expected or appropriate.

Provided $n_A > n_B$, a condition that is met in all data used in the following analysis, option *A* (the "ethical" option) maximizes the sum of monetary payoffs received by the *n* members of a group, minimizes inequality in monetary payoffs, and maximizes the minimum monetary payoff (which in fact holds regardless of whether $n_A > n_B$). Option *B* (the "selfish" option) maximizes the monetary payoff to eligible voters. Note also that the cost of voting ($c = .10$) outweighs the maximum expected monetary benefit ($\frac{x}{2} = .125$) from voting, unless $n_\beta = 1$. If, for example, $n_\beta = 2$, then the expected monetary benefit is $\frac{.125}{2}$.

We conducted a total of six sessions of the experiment in computer labs at Northwestern University (four sessions) and the Experimental Social Science Laboratory (Xlab) at the University of California–Berkeley (two sessions). Subjects were Northwestern or Berkeley undergraduates recruited from the Management and Organizations subject pool, undergraduate social science classes, computer labs

[13] Note that under option *A*, *A* types are paid the same amount $(1 - c)$ as *B* types who vote. The purpose of this (rather than, e.g., simply setting 1 as the payoff to *A* types under option *A*) is to ensure that *B* types who vote for *A* are not compelled to reduce their own payoff *below* that of the beneficiaries of that choice. If individuals dislike payoff disadvantages relative to other subjects, this alternative could confound voters' evaluation of the ethical implications of options *A* and *B*, which we sought to avoid (cf. Charness and Rabin 2002, who found subjects most willing to sacrifice to increase another player's payoff when it was less than their own).

TABLE 2. Experiment Design

	n_A									
n_B	2	3	4	5	6	7	8	9	13	17
1	—	1(10)	—	1(25)	—	—	—	—	—	—
2	2(10)	1(25) 2(10)	—	—	—	—	—	—	—	—
3	2(15)	—	1(10) 2(15) 3(35)	1(10) 2(10) 3(10)	2(15) 3(15)	—	1(15) 3(10)	—	—	—
4	1(10) 4(10)	—	—	1(40) 2(15) 3(25) 4(35)	—	—	—	—	—	1(25) 2(10) 3(10) 4(10)
5	—	—	—	—	2(10) 4(10) 5(15)	1(10) 2(10) 4(10) 5(10)	—	—	—	—
6	—	—	—	—	—	—	—	3(15)	—	—
8	—	—	—	—	—	—	—	1(15) 3(15) 4(10) 6(20) 7(15) 8(10)	1(15) 2(15) 4(10) 7(15) 8(10)	—
11	—	—	—	—	—	—	—	—	2(15) 11(20)	

Entries list number of active *B* types in group, for each possible combination of *A* and *B* types (number of rounds for which the configuration was used in parentheses).

(Northwestern), and the Xlab subject pool (Berkeley). Subjects were not selected to have any specialized training in game theory, political science, or economics. One hundred and four subjects participated across the six sessions, with 18, 11, 9, 24, 21, and 21 subjects, respectively. Each session began with an instruction period to familiarize the participants with the decision problem (instruction script is available at http://polisci.berkeley.edu/faculty/gailmard/mb-instruct.pdf or from the authors upon request), computer software, random matching, and sequence of decisions. Subjects were informed that their groups would be randomly redrawn every round; this was demonstrated in the instruction period by identifying group members in one practice round, and then indicating their different groups in the next practice round. The computer software displayed the payoff table (Table 1) with the experimental parameters, information about the subject's role and the number of subjects in each role in the group in a given round, and the entire history of the subject's own results. All decisions were made in private at computer terminals not visible to other subjects, and all interaction among subjects took place anonymously at computers.

The sessions lasted approximately 100 minutes at Northwestern and 120 minutes at Berkeley, consisted of 90 to 170 rounds, and contained six to fourteen 10 to 20-round blocks of treatments (n_A, n_B, n_β).[14] Subjects were informed of the duration of the session in minutes during recruiting and again in the informed consent process in each session. The number of rounds in each session varied with the duration and the rate of subject decision making; we conducted the maximum number of rounds consistent with the time constraint. The sequence of treatments and rounds used for each treatment block, by session, are listed in Appendix B.

For each subject, five rounds were selected at random at the end of the experiment, and the subject was paid the sum total of his or her earnings in dollars from those rounds, times .04. Participants earned about $25 on average for their session (including a $5 participation payment), with a minimum payment of $5 up to a maximum of about $50. Subjects were paid privately at the end of the session so that a subject and the experimenter knew that subject's payment.

Table 2 summarizes the treatments run over all sessions of the experiment, irrespective of the order or session in which they were run. The rows list n_A values used and the columns list n_B. The cell entries list the values of n_β that were used for each (n_A, n_B) combination (number of rounds in which that value was used in parentheses). Recall that $n = n_A + n_B$ is the number of participants in each group.

Therefore, the possible values of n_β were 1, 2, 3, 4, 5, 6, 7, 8, and 11. Note that in almost all rounds, groups had

[14] A software glitch in session 2 occurred after 18 rounds of the experiment. Eighteen rounds took place before the glitch, 15 in one treatment and 3 in another. Only 10 of 11 subjects were used in these rounds. In total, therefore, session 2 had 85 rounds with 11 subjects and 18 rounds with 10 subjects for a total of 103 rounds.

more A voters than B voters ($n_A > n_B$).[15] Treatments were chosen primarily to maximize the range of possible values of n_β (and therefore pivot probability) given the number of subjects in each session and $n_A > n_B$, while still varying the ratio of n_A to n_B. A consequence is that the design is incomplete in a factorial sense. Note that for most (n_A, n_B) pairs, n_β ranges roughly as much as possible with high contrast between treatments. Treatment choices were constrained by the number of subjects in each session, the requirements to test multiple n_β values while making pivot probability independent of the round of the session,[16] and our desire to limit possible repeated game effects by using (n_A, n_B) combinations that typically allowed more than one group to which subjects could be assigned. Given these constraints, treatments were ordered so that similar (n_A, n_B, n_β) combinations could be tested multiple times with another treatment in between (to help mitigate order effects), and to the extent possible so that similar (n_A, n_B, n_β) combinations could be tested in multiple sessions (to help mitigate session effects).[17]

RESULTS

Individual Behavior and Expressive Voting

Aggregating over all values of pivot probability, ethical voting (option A) occurs with similar frequency as selfish voting (option B), and abstention occurs more frequently than either. In all 2,826 vote choices, 19.5% are A votes, 21% are B votes, and 59.5% are abstentions; conditional on turnout (1,141 observations), 51.9% of all votes are for B. Individual subjects made an average of 30 decisions each, with a range of 10 to 56. A small number of subjects (10) cast an A vote in more than 50% of their decisions; 12 subjects cast a B vote in more than 50% of their decisions. In addition, 44 subjects voted for A strictly more often than B, and 47 voted for B strictly more often. Conditional on turnout, 43 subjects cast an A vote more than 50% of the time, 46 cast a B vote more than 50% of the time, and 12 cast each vote 50% of the time (3 subjects always abstained). Most subjects split their decisions between abstention and either A votes or B votes; few subjects alternated significantly among all three choices. Each subject made his or her least common decision an average of only two times, and for only three subjects was abstention the least common choice. Each subject made his or her most common choice 21.4 times on average, and 76 subjects abstained as often as they made either other choice.[18]

The key implications of the theory and its contrast to purely instrumental voting relate to the effects of pivot probability, and we focus on this for the remainder of the analysis. Our approach and choice of statistical models is aimed at highlighting the effects of pivot probability on vote choices as directly as possible. We use a variety of estimators, model specifications, and levels of analysis to ensure that the conclusions are not excessively dependent on any one choice.

Figure 1 presents graphical evidence on the effect of pivot probability. It shows the percentage of times each choice—ethical voting (A), selfish voting (B), or abstention—is made as a function of pivot probability. Each panel in Figure 1 displays results for a single (n_A, n_B) pair in which more than one value of pivot probability was implemented in the experiment. Hypothesis 1 implies that the dashed "selfish voting" line should slope up (i.e., as pivot probability increases, relatively more selfish votes should be observed). Hypotheses 2 implies that the solid "ethical voting" line should be flat or downward sloping (i.e., as pivot probability increases, relatively fewer ethical votes should be observed).[19] Four of the twelve panels show a uniform decrease in the probability of an ethical vote as pivot probability increases ($(n_A, n_B) =$ (3, 2), (5, 3), (6, 3), (13, 11)). Seven of the twelve panels show a uniform increase in the probability of a selfish vote as pivot probability increases ((n_A, n_B) = (3, 2), (4, 3), (5, 3), (6, 3), (8, 3), (13,11), (17,4)). Some panels reflect results clearly inconsistent with the theory. For instance, in the (9, 8) panel with pivot probability of .125 (a low value in our design), not a single person in 15 trials voted for the ethical alternative. In addition, in the (7, 5) panel, the solid "selfish" votes line is nearly constant (rather than increasing), whereas in the (4, 3) and (8, 3) panels, the share of ethical votes increases with pivot probability.

These interpretations of Figure 1 risk allowing the inevitable variation and noise in the data across rounds and individuals to obscure broader patterns that, although not uniformly true for all values of pivot probability, nevertheless tend to hold. To uncover these general tendencies, we estimated linear probability (OLS) models of the probability of ethical and selfish votes as

[15] We do not use rounds with $n_B \geq n_A$ in any analysis. They were included in the sessions as brief pilot tests for experiments unrelated to the main point of this article.

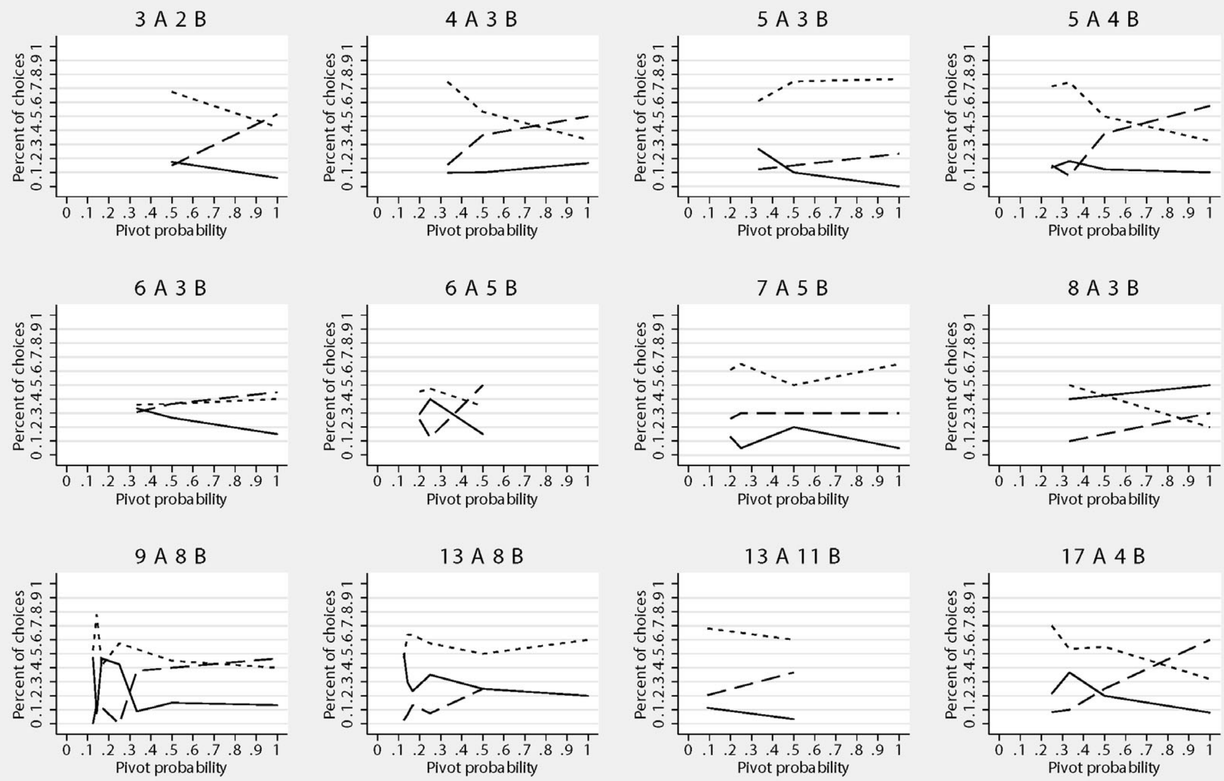
[16] The correlation between pivot probability and round is −.04.

[17] Exact replication of several (n_A, n_B, n_β) combinations suggests that order effects are not a serious problem. In particular, treatments (5,4,4), (13,11,11), (4,3,3), and (17,4,1) were run multiple times in a single session. In none of these cases is the share of selfish votes significantly affected by the order. In only the (17,4,1) treatment was the share of ethical votes significantly lower in the later block ($p = .076$ (2 tailed) in difference in proportions Z test); in other treatments, there is no significant difference between blocks. Note that in eight tests where the null hypothesis of no difference is true, the expected number of false positives for $\alpha = .10$ is .8. For both ethical and selfish voting, there is no consistent pattern of increase or decrease in later blocks. However, there may be some session effects: of four exact (n_A, n_B, n_β) combinations replicated across sessions (3,2,1), (5,1,1), (5,4,1), (5,4,4), two exhibited significant differences between sessions in the share of ethical votes ((5,4,1) and (5,4,4)) and one in the share of selfish votes (5,1,1).

[18] Other summary statistics, including statistics for all variables used in the following statistical models, are in Appendix C.

[19] The theory predicts only a weak inequality for ethical voting because when the expressive utility d is large enough, voters vote ethically regardless of pivot probability, whereas if d is in an intermediate range ethical expressive voting is sensitive to pivot probability because individuals attach relatively more utility to the instrumental costs of ethical voting. The slope of the dotted abstention line is not restricted by the theory.

FIGURE 1. Pivot Probability and Vote Choice by Treatment



Solid line: ethical (A). Dashed line: selfish (B). Dotted line: abstain.

a function of pivot probability in each (n_A, n_B) combination in Figure 1. The resulting coefficients are listed in Table 3 (standard errors clustered by subject[20] are in parentheses; bold entries are significant at the .10 level (2-tailed test)).[21] The final row lists the average coefficient for each vote type.

As Table 3 suggests, the probability of an ethical vote tends to decrease with pivot probability in most treatments, and the probability of a selfish vote tends to increase. For ethical voting, in six of twelve (n_A, n_B) pairs pivot probability has a statistically significant (negative) effect. For selfish voting, in eight of twelve (n_A, n_B) pairs pivot probability has a significantly positive effect. Moreover, for ten of twelve ethical voting regressions, the estimated coefficient is negative; if all slopes were actually zero so the estimated signs were due to chance, at least 10 of 12 would be negative in only 1.9% of samples. All twelve of the selfish voting regressions, have positive estimated effects of pivot probability on the probability of a selfish vote, although, as

TABLE 3. Linear Probability Model Coefficients for Effect of Pivot Probability on Share of Each Vote Type by Treatment

(n_A, n_B)	Ethical	Selfish	N Obs. (*N* clusters)
3, 2	−.23 (.11)	.7 (.17)	146 (26)
4, 3	.09 (.15)	.59 (.19)	435 (42)
5, 3	−.40 (.10)	.17 (.14)	160 (44)
5, 4	−.07 (.11)	.66 (.14)	375 (59)
6, 3	−.19 (.10)	.23 (.29)	105 (27)
6, 5	−.51 (.51)	.87 (.49)	135 (31)
7, 5	−.03 (.06)	.04 (.11)	220 (40)
8, 3	.15 (.18)	.30 (.20)	40 (11)
9, 8	−.29 (.10)	.58 (.17)	365 (60)
13, 8	−.20 (.13)	.21 (.11)	350 (42)
13, 11	−.20 (.06)	.40 (.23)	250 (24)
17, 4	−.23 (.12)	.70 (.17)	135 (40)
Avg Coeff.	−.10	.45	

Clustered SEs in parentheses.

[20] Clustering allows for possible correlation within an individual subject's sequence of votes.

[21] These estimates are not efficient because they neglect the negative correlation in error terms in regressions for a given (n_A, n_B) pair (which occurs because the probability of some vote conditional on pivot probability is 1). Hypothesis test results from Seemingly Unrelated Regressions are essentially the same as reported in Table 3. The one difference is that the slope for ethical voting in the (6, 5) case is significantly negative under the SUR but not the OLS standard errors.

can be seen in Figure 1, the slope for (7,5) is near zero.[22]

Figure 1 is useful because it shows raw data unrestricted by parametric assumptions and effects of pivot probability conditional on group composition. We can go further by aggregating observations across (n_A, n_B) pairs and adding a simple parametric structure to vote choice. We statistically model vote choice as a function of pivot probability,[23] group composition variables n_A and n_B, and the round in which a choice was made. The variables n_A and n_B allow us to probe the effects of group context on vote choice, and although they tend to move together (as Table 2 suggests) their correlation is .50, implying some independent variation by which we can identify the effect of each.[24] The Round variable helps capture experience and some repeated game effects on vote choice.

The following table presents multinomial logit[25] results from the data aggregated from all sessions. Table entries are estimated β parameters, with standard errors clustered by subject to reflect the fact that observations from a particular individual cannot be assumed to be independent. The baseline category is abstention. Therefore, coefficients for option A (the "ethical" option) reflect the effect of each variable on the probability of voting for A as opposed to not voting, and coefficients for option B (the "selfish" option) reflect the effect of each variable on voting for B as opposed to not voting.[26]

Consistent with the theory of ethical expressive voting, the results reveal different effects of pivot probability on the probability of voting for each alternative rather than abstaining ($\chi^2_1 = 22.38$ in a Wald test that these coefficients are equal; $p < .0001$). As implied by Hypothesis 1, the probability that an agent votes selfishly for B rather than abstains is significantly affected by pivot probability: as the pivot probability increases, the probability of voting for B increases. This can be seen by the estimated coefficient on pivot probability in Table 1, which is statistically significant with a Z score of 6.37.[27] The marginal effect of pivot probability (computed at the average) on the probability of voting for B is .33.

In contrast, and consistent with Hypothesis 2, pivot probability has a statistically insignificant effect on the probability a voter votes for A; the Z score for this variable is .45 with a p value of .654.[28] The propensity to vote for A does not decline significantly when pivot probability declines. The estimated marginal effect at the mean is $-.05$, much smaller in absolute value than on voting for B rather than abstaining. This supports the claim that voting for the ethical option is driven at least in part by expressive rather than instrumental factors.

Beyond the implications for the hypotheses, the model also reveals interesting effects of group composition on vote choice. Depending on the ethical considerations voters bring to bear on the choice, both n_A and n_B may affect the ethical weight of A over B. As the number of A types grows (holding fixed the number of Bs), some ethical considerations (e.g., utilitarian or egalitarian ones) point more strongly in favor of alternative A. As the number of B types grows, the pull of these ethical considerations in favor of option A weakens.[29] The results indicate that the decision to vote for the ethical alternative A is significantly affected by the number of A members in the group ($p = .007$) but not the number of B members ($p = .588$). Thus, ethical voting is significantly affected by the number of beneficiaries of that choice but not the number of group members that bear the cost of it. One possible reason for this asymmetry is that voters do not feel an ethical duty to consider the interests of B types because they are not categorically disenfranchised in the experiment; however, the estimated marginal effects are neither large nor significantly different (.009 for A types, $-.007$ for Bs; $\chi^2_1 = 1.69$, $p = .19$ in a Wald test of equal coefficients in the equation for A voting). In

[22] When the number of active Bs is used as the covariate instead of pivot probability, two of the panels with significant ethical voting effects in Table 3 become insignificant ((6,3) and (17,4)). Significance results for selfish voting are the same as in Table 3. The same 10 of 12 ethical voting models have the correct sign on the estimated effect, and all 12 selfish voting models have the correct sign (but with a weakly positive effect on ethical voting and a negative effect on selfish voting).

[23] In none of these models does replacing pivot probability with the number of active B types alter any statistical test results or correspondence between estimated signs and predicted signs of each coefficient.

[24] n_A and n_B have even lower correlation with pivot probability, $-.08$ and $-.45$, respectively.

[25] Hausman and Small-Hsiao tests of the Independence of Irrelevant Alternatives assumption cannot reject the null hypothesis that IIA is satisfied. Essentially, this reflects that no two choices are perceived as close substitutes for each other. In any case, multinomial probit results (which do not depend on IIA) reflect similar effects.

[26] Naturally, there are many ways to estimate comparative statics at the individual level, and our results are not particularly dependent on the multinomial logit model. For instance, seemingly unrelated regressions for ethical and selfish votes with fixed effects for individual subjects, estimated with the same covariates listed here, reveal similar substantive and statistical results. Probably the most important difference due to this alternative estimator is that the estimated effect of pivot probability on ethical votes is negative and more significant than in the multinomial logit specification, although $p > .10$. Note that random assignment of individuals to treatments suggests that individual tendencies to vote in a particular way should not be systematically related to other covariates, so accounting for individual heterogeneity in estimation is not crucial as it is in observational data.

[27] It might be suspected that the number of B types is highly correlated with pivot probability, because the latter is isomorphic to the number of active B types. However, this correlation is about $-.44$. In any case, multicollinearity problems would typically be suspected of making intercorrelated variables appear falsely insignificant, not falsely significant.

[28] A different way to see the same information contained in the p value is a *post hoc* power calculation, that is, the probability of rejecting $H_0 : \beta_{piv.prob} = 0$ at the $\alpha = .10$ level if the true coefficient $\beta_{piv.prob}$ and sample variability are in fact as estimated in the model. Under these assumptions, the lower bound of the rejection region for H_0 is $\beta = .645$, and if the true parameter is .175 with standard deviation .391 in its sampling distribution, the estimated effect is (falsely) in the rejection region with probability is .115 under H_0.

[29] Given fixed payoffs for options A and B, only efficiency considerations are affected by group composition in this way. For example, maximin considerations are unaffected by the size and composition of each group.

contrast, selfish voting is not significantly affected by the number of A or B types. If A is the ethical alternative, then according to our theory votes for B are due to selfish considerations; the results are consistent with this, in that any effects of group composition on the ethical pull of A over B do not significantly affect votes for the selfish alternative.[30] Although the precise ethical considerations voters use in making decisions is beyond the scope of our analysis, these results suggest that some natural interpretations of the ethical implications of group composition do affect voting decisions.

The results show some effects of experience on both ethical and selfish voting, although the p values are at most marginally significant at .19 and .14, respectively. This is reflected in the Round variable, which indexes the round of the session in which a decision occurred. Moreover, the substantive effects of experience are small, about $\frac{1}{20}$ of 1% for both ethical and selfish voting. Thus, over an entire session, the probability of making either vote choice changes by a small amount, conditional on group composition and pivot probability. These results help alleviate concerns about both experience and repeated game effects in our design.[31]

The multinomial results can be sharpened by exploring treatment effects on vote choice, conditional on turnout. In the previous model, the expressive component of utility exerts a greater effect on decisions *of a given voter* as pivot probability declines (the preference effect). This in turn implies that, conditional on turnout, voting for the ethical option grows more likely as pivot probability declines (Hypothesis 3). To test this hypothesis, we model vote choice conditional on turnout, as a function of the same covariates specified previously. In this case, we estimate a logit model with subject fixed effects, so that only variation within each subject's sequence of votes is used to estimate parameters, and variation across subjects is discarded.[32] The reason is that individuals self-select into the voting population, so that random assignment of subjects to treatments does not ensure that the effect of pivot probability *conditional on turnout* is unrelated to unobserved individual characteristics (in particular, the strength of individuals' ethical preferences). Put differently, given our design, within-subjects information is the best way to identify the preference effect underlying Hypothesis 3, which according to the theory operates on each individual voter. Between-subjects information is confounded by self-selection into the voting population.

TABLE 4. Individual Vote Choice (Multinomial Logit; Base Category Is Abstain)

Covariate	Parameter Estimate	Clustered SE
Effect on Pr (Vote for *A*)		
Number of A types	.048***	.018
Number of B types	−.025	.046
Pivot probability	.175	.391
Round	.002	.002
Constant	−1.612***	.332
Effect on Pr (Vote for *B*)		
Number of *A* types	−0.033	.022
Number of *B* types	.083	.054
Pivot probability	2.100***	.330
Round	−.003	.002
Constant	−1.914***	.344

2,826 observations; standard errors clustered for 104 subjects.
Note: * Indicates $p < .10$, ** indicates $p < .05$, *** indicates $p < .01$.

Consistent with Hypothesis 3 and the preference effect, a decline in pivot probability does make a vote for the ethical alternative more likely conditional on turnout. The result is highly significant with a Z score of 5.96. Again, consistent with the operation of the preference effect at the level of the individual voter, this result stems from changes within subjects' sequences of votes as pivot probability changes.

Similar to the decision to vote ethically rather than abstain, the decision to vote selfishly rather than ethically is negatively affected by n_A ($p = .001$) but not n_B (although this achieves marginal significance, $p = .126$). Unlike the multinomial results, the effects of these variables are significantly different in absolute value ($\chi_1^2 = 5.52, p = .019$ in a Wald test of equal coefficients), so the asymmetric effect of n_A and n_B on ethical voting carries over to the size as well as significance. The effect of Round shows that agents are significantly less likely to vote selfishly, conditional on turnout, in later rounds than earlier ones. Combined with Table 4, this suggests mixed evidence on experience effects, but even the significant estimate in Table 5 does not suggest repeated game effects; subjects are estimated to become *less* likely to vote selfishly in later rounds conditional on group composition and pivot probability. One possible reason for this experience effect is that subjects take time to discover the ethical implications of each option, because the instructions and experiment software purposely made no mention of any ethical considerations when describing the choice environment.

Individual Behavior and Instrumental Voting

Levine and Palfrey (2007) use a laboratory experiment to demonstrate that quantal response equilibrium

[30] It should be noted that although the effects of n_A and n_B on selfish voting do not meet conventional significance levels, they are marginally significant with p values of .15 and .13, respectively. Nevertheless, the effect of n_A on ethical voting is significantly greater than for selfish voting ($\chi_1^2 = 9.09, p = .003$ in a Wald test of equal coefficients across equations), whereas the effect of n_B is not significantly different across equations ($\chi_1^2 = 1.97, p = .160$).

[31] Repeated game effects imply that the future costs of selfish behavior (if it induces others to act selfishly as well) decline toward the end of a session, and thus the small and insignificant effect of the Round variable casts doubt on this explanation of ethical voting. In addition, because smaller sessions imply a larger number of interactions with any given subject, we estimated the multinomial logit model with the number of subjects in the session as a covariate. Its effect is statistically insignificant for both A and B votes (and the significance of pivot probability is unchanged from Table 4). Beyond these points, it is unclear why repeated play would increase votes for A disproportionately when pivot probability is low.

[32] We also estimated, and obtained similar results from, a linear probability model with subject fixed effects.

TABLE 5. Vote Choice Conditional on Turnout (Fixed Effects Logit)

Covariate	Parameter Estimate	Standard Error
Effect on Pr (Vote for *B*)		
Number of *A* types	−.115***	.035
Number of *B* types	.104	.068
Pivot probability	3.287***	.454
Round	−.010***	.003

674 observations; fixed effects for 59 subjects. The remaining subjects either always voted for *A* or always voted for *B* conditional on turnout.
Note: * Indicates $p < .10$, ** indicates $p < .05$, *** indicates $p < .01$.

models with selfish instrumental voters can explain turnout and vote choice in large elections. Our experiment presents subjects with a decision situation, not a game, so in our unified model, quantal response equilibrium (McKelvey and Palfrey 1995) collapses to quantal choice or random utility maximization (Luce 1959; McFadden 1974), and a logit QRE is simply a logit random utility model. Because individual behavior is clearly not deterministic conditional on observable predictors, some form of random utility model is necessary to explain the results regardless of whether ethical expressive utility is assumed. Thus, any departure from selfish instrumental voting in our data is not related to the quantal choice element but rather to the arguments of individual utility functions.

To see the effects of selfish instrumental preferences, assume that d and δ from the formal model are zero so that behavior conforms to a selfish quantal choice model. Three implications of this special case emerge:

1. The probability of voting for each alternative rather than abstaining declines as pivot probability declines. This is because, given that voting is costly (and dominated if $n_\beta > 2$), the error of turning out becomes more costly as pivot probability declines.
2. Conditional on voting and pivot probability, the probability an agent votes for B is always greater than the probability he or she votes for A. This is because, in quantal choice models, the probability of playing better strategies exceeds that of playing worse strategies.
3. Conditional on voting, the probability of voting for A increases and the probability of voting for B decreases as pivot probability declines. This is because the error of voting for A becomes less costly as pivot probability declines, because it is less likely to affect the outcome.

Implication 1 is not supported in our data, as noted in the multinomial choice results: only selfish voting is depressed by increases in pivot probability; ethical voting is not. Implication 2 is also not supported in our data: as Figure 1 demonstrates, for relatively small values of pivot probability, the share of A choices is sometimes larger than the share of B choices. Specifically, the share of A votes exceeds the share of B votes in 18 of the 46 distinct (n_A, n_B, n_β) triples in our data. Implication 3 follows from both the selfish instrumental and ethical expressive models, when a quantal choice component is included in each. It is supported in our data in the sense that agents are more likely to vote for A as pivot probabilities decline.

Overall, then, selfish instrumental behavior alone cannot account for our findings even if a quantal choice element is included. Rather, our findings suggest that ethical expressive motivations play an important role in voting, but not to the exclusion of selfishness (i.e., the ethical expressive conjecture does not deny the role of selfish instrumental preferences; rather, it layers another motivation on top of this). In one sense, it is not surprising that our results differ from Levine and Palfrey's (2007) in the extent to which selfishness can account for the findings. In their design, members of one group of voters obtain almost all benefits if a first alternative wins (105 points compared to 5 points); the other group of voters obtains almost all benefits if the second alternative wins. Even if members of the second group believe the first choice is ethically desirable, it is materially costly for them to act on this in casting a vote. In our design, voters sacrifice smaller material benefits to vote ethically. Moreover, although the groups in Levine and Palfrey's design have different sizes, so the alternatives have different utilitarian implications, neither alternative is unambiguously ethical in their design. Cross-cutting ethical considerations might mitigate their effect (by lowering the difference in expressive utility between alternatives), leaving selfish voting with greater weight in voting decisions.

An alternative theory of instrumental voting specifies that $\delta > d = 0$, so that individuals value the ethical outcome but not the act of voting for that outcome per se. Such accounts have been forwarded recently by Jankowski (2002); Fowler (2006); and Edlin, Gelman, and Kaplan (2007). Inasmuch as ethical voting does occur, even with pivot probability equal to 1 we see some evidence of $\delta > 0$, but this purely instrumental account of ethical behavior is not an adequate explanation for our findings. First, conditional on group size, decreasing pivot probability should still depress participation rates for these instrumental A voters. Yet, the results show that abstention is not significantly more likely relative to voting for A as pivot probability declines, holding fixed the number of A and B types. A possible counterargument to this evidence is that for these ethical instrumental voters, the altruistic benefits of voting are so high that even our lowest pivot probability does not depress their turnout. This leads to a second problem with this explanation for our findings. Linear utilities with constant altruism weights map straightforwardly into a structural model that is estimable in our data. Multinomial logit estimates show that for this model to explain our experimental data, active B types must value an A type's payoff as much as they value their own, and must value payoffs to other B types four times as much as they value their own. These implausible findings suggest that a model consisting only of ethical instrumental voters does not fare well in our data.

TABLE 6. Effect of Group Characteristics on Group Choice (Logit)

Covariate	Parameter Estimate	Het. Robust SE
Effect on Pr (Group choice is *B*)		
A types	−.022	.021
B types	.071*	.039
Pivot probability	.550**	.221
Random choice	−.307**	.130
Round	.002	.002
Constant	−.138	.262

1,061 observations; heteroskedasticity-robust standard errors.
Note: * Indicates $p < .10$, ** indicates $p < .05$, *** indicates $p < .01$.

Collective Choices and Bias toward Unselfishness

The previous results show the importance of ethical expressive considerations for understanding individual behavior in our experiment. But this does not by itself demonstrate the importance of these considerations for understanding collective choices. The individual-level effects may be small or ethical expressive motivations so unusual as to be irrelevant for understanding the functioning of elections in collective decision making. Moreover, group decisions are made by individuals who are pivotal, and obviously, a disproportionate share of pivotal votes were cast in situations where pivot probability is large. Because the ultimate importance of our analysis lies in how political institutions translate preferences into collective decisions in the presence of nonselfish, expressive agents, the question of group-level behavior is also important, and we turn to it now.

Table 6 presents logit results on the effect of group characteristics on the probability that the group choice was the selfish option B rather than the ethical option A.[33] Because the variance of the observed outcome could change with the group characteristics, we report heteroskedasticity-robust standard errors. The model is estimated based on all group decisions, including those made at random rather than by an active B type (summary statistics of these variables are in Appendix C). These observations are important to include because they create a possible disconnect between the individual-level results and group decisions. Relevance of ethical expressive voting at the group level requires that these random decisions do not overwhelm the effect of pivot probability on individual turnout and vote choice decisions.

The most important feature of the results is that they support Hypothesis 4. As pivot probability increases, the probability of a selfish choice for the group increases. The coefficient on pivot probability has a p value of .013, and the marginal effect (at its average, 57.4% in this sample) is .14. In short, the individual-level effects of pivot probability identified previously carry over to group choices as well.

Turning to the other covariates, the effects of group composition are reversed compared to the individual level. The effect of the number of B types is statistically significant ($p = .067$), but the effect of the number of A types is not ($p = .294$). Moreover, the effect of n_B is significantly larger than the effect of n_A in absolute value ($\chi^2_1 = 2.92$, $p = .088$ in Wald test of equal coefficients). The negative effect of Random Choice on the probability of B reflects that the selfish choice was significantly more likely to be implemented by a subject in the experiment (roughly 61% of these cases) than by a random draw when the selected voter abstained (roughly 52% of these cases). Group-level choices are not significantly affected by the experience of participants; p value for Round is .26.

CONCLUSION

In this article, we provide experimental support for an ethical expressive model of voting. In our experiment, groups must choose between two options—an "ethical" option with a relatively equal distribution of payoffs that maximizes total payoffs and the minimum payoff in the group, and a "selfish" option with a lopsided distribution favoring the voters themselves. Our design allows us to manipulate the distribution of payoffs from each option and, most important, the pivot probability of individual voters.

The experimental results support the concept of bias toward unselfish outcomes in large elections: collective choices in elections systematically depart from individual preferences in the direction of moral considerations as pivot probabilities decline. The data are consistent with a significant effect of ethical expressive preferences. Bias toward unselfish outcomes then results for two distinct reasons: first, as the pivot probability declines, the choice of any agent that actually votes becomes closer to a hypothetical choice, in which case ethical considerations dominate selfish ones (the preference effect). Second, as pivot probability declines, voters with relatively strong instrumental and relatively weak expressive preferences are less likely to vote, whereas voters with relatively strong expressive preferences may continue to vote or switch from abstention to voting for the ethical option. Thus, the ratio of ethical expressive voters to instrumental voters grows as pivot probability declines (the turnout effect).

APPENDIX A: BENCHMARK VOTING MODELS

In this appendix, we demonstrate that our specifications of payoffs in the model can be seen as integrating three different polar types of voter: selfish instrumental, ethical instrumental, and ethical expressive. These are special cases of the general formal model presented previously.

Selfish Instrumental Motivations

The standard model of selfish instrumental voting results of $d = \delta = 0$, and there are no payoff disturbances. In particular,

[33] Results are similar in a linear probability model as well as a logit model with fixed effects for (n_A, n_B) pairs (although in the latter case, the effects of n_A and n_B obviously cannot be estimated).

the payoff to voting for B is

$$\frac{1}{n_\beta}(1+x)+\left(1-\frac{1}{n_\beta}\right)(1+q^*x)-c,$$

whereas before $\frac{1}{n_\beta}$ is the pivot probability and q^* is the probability option B is chosen when his or her vote is not pivotal. The payoff from abstaining is

$$\frac{1}{n_\beta}\left(1+\frac{x}{2}\right)+\left(1-\frac{1}{n_\beta}\right)(1+q^*x).$$

Thus, the selfish B type weakly prefers to vote for B rather than abstain if and only if

$$\frac{x}{2n_\beta}\geq c.$$

So, as the probability a vote is pivotal ($\frac{1}{n_\beta}$) decreases, the incentive for a selfish B type to abstain gets larger. For given x and c, there is a value of pivot probability such that voting is a dominated strategy for any smaller value of pivot probability. In addition, conditional on choosing to vote, a selfish B type has a strictly dominant strategy to vote for B.

Ethical Instrumental Motivations

An alternative model results when voters obtain a payoff δ and option A is chosen. However, if $\delta > x$, the decision to vote (as oppose to abstain) is entirely analogous to the selfish instrumental case. The payoff to voting for option A is

$$\frac{1}{n_\beta}(1+\delta)+\left(1-\frac{1}{n_\beta}\right)(1+q^*x+(1-q^*)\delta)-c,$$

whereas the payoff for abstaining is

$$\frac{1}{n_\beta}\left(1+\frac{x}{2}+\frac{\delta}{2}\right)+\left(1-\frac{1}{n_\beta}\right)(1+q^*x+(1-q^*)\delta).$$

Hence, with ethical instrumental payoffs, voting for A is beneficial if and only if

$$\frac{\delta-x}{2n_\beta}\geq c.$$

So, as the probability of being pivotal decreases (n_β increases), the incentive to participate decreases.

Ethical Expressive Motivations

A third polar case occurs when voters obtain an expressive payoff $d > c$ by voting for A. The payoff to voting for option A is

$$\frac{1}{n_\beta}+\left(1-\frac{1}{n_\beta}\right)(1+q^*x)+d-c,$$

whereas the payoff for voting for option B is

$$\frac{1}{n_\beta}(1+x)+\left(1-\frac{1}{n_\beta}\right)(1+q^*x)-c$$

Conditional on voting, voters prefer to vote for A over B if

$$d\geq\frac{x}{n_\beta}.$$

So, conditional on voting, as the probability of being pivotal decreases the incentive to vote for A increases. Note that this is in contrast to both the selfish and ethical instrumental benchmark cases where pivot probabilities do not impact the choice between A and B.

When $d\geq\frac{x}{n_\beta}$, a voter prefers to vote for A rather than abstain if and only if

$$\frac{1}{n_\beta}(1)+\left(1-\frac{1}{n_\beta}\right)(1+q^*x)+d-c \geq \frac{1}{n_\beta}\left(1+\frac{x}{2}\right)+\left(1-\frac{1}{n_\beta}\right)(1+q^*x)$$

or

$$d-c\geq\frac{x}{2n_\beta}.$$

When $d<\frac{x}{n_\beta}$, a voter prefers to vote for B rather than abstain if and only if

$$\frac{1}{n_\beta}(1+x)+\left(1-\frac{1}{n_\beta}\right)(1+q^*x)-c \geq \frac{1}{n_\beta}\left(1+\frac{x}{2}\right)+\left(1-\frac{1}{n_\beta}\right)(1+q^*x)$$

or

$$\frac{x}{2n_\beta}\geq c.$$

Consider three different cases: $d\geq x$ (d large), $x>d>2c$ (d intermediate), and $2c>d$ (d low). When d is large, the ethical expressive voter always votes for A. When d is intermediate, this type votes for B when the pivot probability is high ($\frac{1}{n_\beta}>\frac{d}{x}$), votes for A otherwise. When d is low, then this type votes for B when the pivot probability is large ($\frac{1}{n_\beta}>\frac{2c}{x}$), abstains when the pivot probability is in the interval ($\frac{2(d-c)}{x}$, $\frac{2c}{x}$), and votes for A when d is small ($\frac{2(d-c)}{x}>\frac{1}{n_\beta}$).

This optimal behavior is derived under all possible parameter values for c and d as follows.

Case 1: $d>2c$	$n_\beta<\frac{x}{d}$	$n_\beta>\frac{x}{d}$
$x>d$	Vote B	Vote A
$x<d$	—	Vote A

Case 2: $2c>d>c$	$\frac{x}{2c}>n_\beta$	$\frac{x}{2(d-c)}>n_\beta>\frac{x}{2c}$	$n_\beta>\frac{x}{2(d-c)}$
$x>2c$	Vote B	abstain	Vote A
$x<2c$	—	abstain	Vote A

Case 3: $c>d>0$	$\frac{x}{2c}>n_\beta$	$n_\beta>\frac{x}{2c}$
$x>2c$	Vote B	abstain
$x<2c$	—	abstain

The results can be grouped into six cases (one for each row of each table) of the relationship between pivot probability and vote choice. The cases list the sequence of optimal decisions for ethical expressive voters as pivot probability declines below the cutpoints in the tables. In all cases, the incentive to vote B is strongest for high pivot probabilities, the incentive to vote A is strongest for low pivot probabilities, and the incentive to abstain is highest for moderate pivot probabilities.

Case 1 ($d>2c$ and $d<x$): vote B, vote A
Case 1* ($d>2c$ and $d>x$): vote A
Case 2 ($2c>d>c$ and $2c<x$): vote B, abstain, vote A
Case 2* ($2c>d>c$ and $2c>x$): abstain, Vote A
Case 3 ($d<c$ and $x>2c$): vote B, abstain
Case 3* ($d<c$ and $x<2c$): abstain.

Ethical Expressive Voting When Option *B* Is the Ethical Alternative

Ethical expressive types get the same payoffs as selfish voters plus a payoff of $d > 0$ by voting for option A. The payoff to this type for voting for option A is

$$\frac{1}{n_\beta}(1+\delta) + \left(1 - \frac{1}{n_\beta}\right)(1 + q^* x + (1-q^*)\delta) + d - c,$$

whereas the payoff for voting for option B is

$$\frac{1}{n_\beta}(1+x) + \left(1 - \frac{1}{n_\beta}\right)(1 + q^* x + (1-q^*)\delta) - c.$$

Conditional on voting, ethical expressive voters prefer to vote for A over B if

$$d \geq \frac{x-\delta}{n_\beta}.$$

So, conditional on voting, as the probability of being pivotal decreases the incentive for an ethical expressive type to vote for A increases.[34] This is in contrast to both the selfish and ethical instrumental models, where pivot probabilities do not impact the choice between A and B.

Voters with $d \geq \frac{x-\delta}{n_\beta}$ prefer to vote for A rather than abstain if and only if

$$\frac{1}{n_\beta}(1+\delta) + \left(1 - \frac{1}{n_\beta}\right)(1 + q^* x + (1-q^*)\delta) + d - c$$
$$\geq \frac{1}{n_\beta}\left(1 + \frac{x+\delta}{2}\right) + \left(1 - \frac{1}{n_\beta}\right)(1 + q^* x + (1-q^*)\delta)$$

or

$$d - c \geq \frac{x-\delta}{2n_\beta}.$$

Voters with $d < \frac{x-\delta}{n_\beta}$ prefer to vote for B rather than abstain if and only if

$$\frac{1}{n_\beta}(1+x) + \left(1 - \frac{1}{n_\beta}\right)(1 + q^* x + (1-q^*)\delta) - c$$
$$\geq \frac{1}{n_\beta}\left(1 + \frac{x+\delta}{2}\right) + \left(1 - \frac{1}{n_\beta}\right)(1 + q^* x + (1-q^*)\delta)$$

or

$$\frac{x-\delta}{2n_\beta} \geq c.$$

Previous experimental work (Levine and Palfrey 2007) shows substantial support for the selfish voter model in comparative statics. Their results find that turnout in laboratory experiments conforms with the comparative statics predicted by the selfish model (e.g., turnout is decreasing as the size of the electorate increases, and therefore, pivot probabilities decrease). In this section, we show that it is difficult to differentiate between the selfish and ethical expressive model when the alternative that is favored by selfish voters is also perceived to be the ethical alternative. We show that the only difference between the two models is in the level of turnout predicted. In the ethical expressive model, turnout does not go to zero as pivot probabilities get small.

Suppose that ethical expressive types get a payoff of $\delta > 0$ when alternative B is chosen and a payoff of $d > 0$ by voting for option B. It is obvious that such voters will never vote for option A. The payoff for voting for option B is

$$\frac{1}{n_\beta}(1+x+\delta) + \left(1 - \frac{1}{n_\beta}\right)(1 + q^*(x+\delta)) + d - c.$$

Subjects prefer to vote for B rather than abstain if and only if

$$\frac{1}{n_\beta}(1+x+\delta) + \left(1 - \frac{1}{n_\beta}\right)(1 + q^*(x+\delta)) + d - c$$
$$\geq \frac{1}{n_\beta}\left(1 + \frac{x+\delta}{2}\right) + \left(1 - \frac{1}{n_\beta}\right)(1 + q^*(x+\delta))$$

or

$$d - c \geq -\frac{x+\delta}{2n_\beta}.$$

So, if $d - c > 0$ the subject votes, whereas if $d - c < 0$ then the probability of voting is decreasing in n_β. It follows that turnout is decreasing as pivot probabilities decrease but reaches a lower bound. Note that Levine and Palfrey (2007) observe in their experiments that turnout levels in elections with low pivot probabilities seem to be bounded significantly above zero.

APPENDIX B: TREATMENT SEQUENCE BY SESSION

TABLE A1. Treatments by Session

	Session					
Block	1	2	3	4	5	6
1	(9,8,1) 15	(2,3,2) 15	(3,1,1) 10	(5,3,3) 10	(4,3,2) 15	(4,3,3) 15
2	(5,4,1) 15	(3,2,1) 3	(2,2,2) 10	(5,3,1) 10	(13,8,2) 10	(13,8,4) 10
3	(9,8,7) 15	(6,5,5) 15	(5,4,4) 15	(7,5,1) 10	(13,8,6) 15	(6,5,4) 10
4	(5,4,2) 15	(5,6,1) 15	(6,3,2) 15	(7,5,5) 10	(4,3,1) 10	(17,4,1) 10
5	(5,1,1) 20	(8,3,1) 10	(6,3,3) 15	(13,11,11) 10	(9,8,6) 10	(17,4,3) 10
6	(9,8,3) 15	(5,1,1) 10	(5,4,2) 15	(13,11,2) 15	(9,8,2) 10	(6,5,2) 10
7		(5,4,1) 10	(5,4,4) 10	(13,11,11) 10	(4,3,3) 10	(6,3,1) 15
8		(5,4,4) 10		(7,5,4) 10	(17,4,4) 15	(17,4,2) 10
9		(8,3,3) 10		(2,4,1) 10	(13,8,1) 15	(13,8,8) 10
10				(2,4,4) 10	(13,8,7) 15	(4,3,3) 10
11				(18,6,2) 10	(3,2,2) 10	(5,4,3) 10
12					(3,2,1) 15	(17,4,1) 15
13					(7,5,2) 10	(5,3,2) 10
14					(9,8,4) 10	(9,8,8) 10
	18 subs.	11 subs.	9 subs.	24 subs.	21 subs.	21 subs.

Entries are (n_A, n_B, n_β) triples and number of rounds over which each given treatment block was run.

[34] As the pivot probability increases, the set of pairs (d, δ) that satisfy the previous equation decreases (by inclusion).

APPENDIX C: SUMMARY STATISTICS

TABLE A2. Summary Statistics, Individual Decisions

Variable	Mean	SD	Range
Abstentions	59.6%	49.1%	Binary
A votes	19.4%	39.6%	Binary
B votes	20.9%	40.7%	Binary
B votes | turnout	51.9%	50.0%	Binary
Number of *A* types per group	7.9	4.0	[3, 18]
Number of *B* types per group	5.3	2.8	[1, 11]
Pivot probability	37.5%	27.3%	[0.091, 1]
Round	66.8	42.8	[1, 170]

Based on 2,826 observations of individual decisions.

TABLE A3. Summary Statistics, Group Decisions

Variable	Mean	SD	Range
Outcome ($A = 0$ or $B = 1$)	56.5%	49.6%	Binary
Number of *A* types per group	6.8	3.9	[3, 18]
Number of *B* types per group	4.2	2.4	[1, 11]
Pivot probability	57.4%	33.0%	[0.091, 1]
Random choice (1 = yes)	54.8%	49.8%	Binary
Round	66.5	44.4	[1, 170]

Based on 1,061 observations of group outcomes.

REFERENCES

Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow." *Economic Journal* 100: 464–77.

Andreoni, James. 2006. "Philanthropy." In *Handbook of Giving, Reciprocity, and Altruism*, eds. L.-A. Gerard-Veret, S.-C. Kolm, and J. Mercier. New York: Elsevier.

Babcock, Linda, and George Lowenstein. 1997. "Explaining Bargaining Impasse: The Role of Self-Serving Biases." *Journal of Economic Perspectives* 11: 109–26.

Beckman, Steven, John Formby, W. James Smith, and Buhong Zheng. 2002. "Envy, Malice, and Pareto Efficiency: An Experimental Investigation." *Social Choice and Welfare* 19: 349–67.

Bolton, Gary, and Axel Ockenfels. 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90: 166–93.

Brennan, Geoffry, and James Buchanan. 1984. "Voter Choice: Evaluating Political Alternatives." *American Behavioral Scientist* 28: 185–201.

Brennan, Geoffrey, and Loren Lomasky. 1993. *Democracy and Decision: The Pure Theory of Electoral Preference.* New York: Cambridge University Press.

Camerer, Colin. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton, NJ: Princeton University Press.

Carter, John, and Stephen Guerette. 1992. "An Experimental Study of Expressive Voting." *Public Choice* 73: 251–60.

Charness, Gary, and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117: 817–69.

Coate, Stephen, and Michael Conlin. 2004. "A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence." *American Economic Review* 94: 1476–504.

Cummings, R. G., G. Harrison, and E. E. Rutström. 1995. "Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive Compatible?" *American Economic Review* 85: 260–6.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.

Duffy, John, and Margit Tavits. 2006. "Beliefs and Voting Decisions: A Test of the Pivotal Voter Model." Typescript, University of Pittsburgh.

Edlin, Aaron, Andrew Gelman, and Noah Kaplan. 2007. "Voting as a Rational Choice: Why and How People Vote to Improve the Well-being of Others." *Rationality and Society* 19: 293–314.

Feddersen, Timothy, and Wolfgang Pesendorfer. 1996. "The Swing Voter's Curse." *American Economic Review* 86: 408–24.

Feddersen, Timothy, and Alvaro Sandroni. 2006a. "A Theory of Participation in Elections." *American Economic Review* 96: 1271–82.

Feddersen, Timothy, and Alvaro Sandroni. 2006b. "Ethical Voters and Costly Information Acquisition." *Quarterly Journal of Political Science* 1 (3): 287–311.

Fehr, Ernst, and Klaus Schmidt. 1999. "A Theory of Fairness, Competition, and Co-operation." *Quarterly Journal of Economics* 114: 817–68.

Fischer, A. J. 1995. "A Further Experimental Study of Expressive Voting." *Public Choice* 88: 171–80.

Fowler, James 2006. "Altruism and Turnout." *Journal of Politics* 68: 674–83.

Frohlich, Norman, and Joe Oppenheimer. 1992. *Choosing Justice.* Berkeley: University of California Press.

Grether, David, and Charles Plott. 1979. "Economic Theory and the Preference Reversal Phenomenon." *American Economic Review* 69: 623–38.

Harrison, Glenn W., and E. Elisabet Rutström. 2008. "Experimental Evidence on the Existence of Hypothetical Bias in Value Elicitation Methods." In *Handbook of Experimental Economics Results*, eds. Charles Plott and Vernon Smith. New York: North-Holland.

Harsanyi, John. 1977. "Rule Utilitarianism and Decision Theory." *Erkenntnis* 11: 25–53.

Jankowski, Richard. 2002. "Buying a Lottery Ticket to Help the Poor: Altruism, Civic Duty, and Self Interest in the Decision to Vote." *Rationality and Society* 14: 55–77.

Kinder, Donald, and D. Roderick Kiewiet. 1979. "Economic Grievances and Political Behavior: The Role of Personal Discontents and Collective Judgments in Congressional Voting." *American Journal of Political Science* 23: 495–527.

Levine, David, and Thomas Palfrey. 2007. "The Paradox of Voter Participation? A Laboratory Study." *American Political Science Review* 101: 143–58.

List, John, and Craig Gallet. 2001. "What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values? Evidence from a Meta-Analysis." *Environmental and Resource Economics* 20: 241–54.

Luce, Duncan. 1959. *Individual Choice Behavior: A Theoretical Analysis*. New York: Wiley.

McFadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers of Econometrics*, ed. P. Zarembka. Academic Press.

McKelvey, Richard, and Thomas R. Palfrey. 1995. "Quantal Response Equilibrium in Normal Form Games." *Games and Economic Behavior* 10: 6–38.

Messick, David, and Keith Sentis. 1979. "Fairness and Preference." *Journal of Experimental Social Psychology* 15: 418–34.

Morton, Rebecca B. 1991. "Groups in Rational Turnout Models." *American Journal of Political Science* 35: 758–76.

Riker, William, and Peter Ordeshook. 1968. "A Theory of the Calculus of Voting." *American Political Science Review* 62: 25–43.

Scheussler, Alex. 2000. *A Logic of Expressive Choice*. Princeton, NJ: Princeton University Press.
Sears, David O., and Jack Citrin. 1985. *Tax Revolt: Something for Nothing in California*. Cambridge, MA: Harvard University Press.
Sears, David O., and Carolyn L. Funk. 1991. "The Role of Self-Interest in Social and Political Attitudes." *Advances in Experimental Social Psychology* 24: 1–91.
Tullock, Gordon. 1971. "The Charity of the Uncharitable." *Western Economic Journal* 9: 379–92.
Tyran, Jean-Robert. 2005. "Voting When Money and Morals Conflict: An Experimental Test of Expressive Voting." *Journal of Public Economics* 88: 1645–64.
Tyran, Jean-Robert, and Rupert Sausgruber. 2006. "A Little Fairness May Induce a Lot of Redistribution in Democracy." *European Economic Review* 50: 469–85.
Uhlaner, Carol. (1989). "Relational Goods and Participation: Incorporating Sociability into a Theory of Rational Action." *Public Choice* 62: 255–85.